[독립심화학습 결과보고서]

팀 명	김정운	
과 제 명	국문	타겟을 추적하는 UV 에이전트에 대한 최적 경로의 존재성
	영문	The existence of optimal path for agent tracking target

1. 과제 개요(선정 배경 및 취지, 내용 등)

무인 개체(Unmanned Vehicle Agent, 이하 UVUV 에이전트)의 타겟 추적 알고리즘은 오랫동안 연구되었다. 그리고 최근에는 3 차원 유클리드 공간이 아닌 구면에서의 타겟 추적과 관련된 연구도 진행 중이다(Sun-Ho Choi, Dohyun Kwon, Hyowon Seo, 2023, [1]). 해당 논문은 타겟과 이를 추적하는 UV 에이전트 간의 관계를 미분방정식으로 기술했으며, 리아푸노브 함수(lyapunov function)를 사용하여 추적이 가능하다는 것을 증명했다. 이러한 모델링은 상미분방정식으로 UV 에이전트의 위치와 속도를 기술했기 때문에, 이를 기존의 알고리즘보다 빠르게 구할 수 있다. 그러나 해당 논문에서 UV 에이전트를 제어하는 2 개의 control term 은 몇 문제를 가지고 있다.

우선 해당 control term 들이 최적의 경로를 산출한다고 할 수 없는데, 이들은 단순히 리아푸노프 수렴성을 위해 임의로 설정되었기 때문이다(Sun-Ho Choi, Dohyun Kwon, Hyowon Seo, 2023, [1]). UV 에이전트의 이동거리, 임무 수행 시간 및 에너지를 줄이는 등, 다양한 방법으로 UV 에이전트를 보다 효율적으로 운용할 수 있다. 하지만 논문의 control term 은 그러한 비용을 최소화한다는 것을 보장하지 않는다. 또한 [1]의 control 중 하나는 상수 0 을 취하는데, 이는 타겟의 정보를 활용하지 않기 때문에 비효율적이다. 마지막으로 현실의 UV 에이전트가 control 을 실행하지 못할 수 있는데, 개체의 이동능력과 관련된 부등식이 설정되지 않았기 때문이다.

이러한 문제점을 해결하기 위해 UV 에이전트에 대한 제한조건과 비용함수를 설정하고, 이를 최적제어문제(optimal control problem)를 통해 control 을 얻고자 한다. 문제를 단순하게 설정하기 위해, 타겟의 위치는 고정되었으며 3 차원 유클리드 공간을 가정한다. 하지만 갈릴레오 변환을 사용하면, 본 연구를 움직이는 타겟에도 적용할 수 있다, 글의 순서는 "optimal control problem, Pontryagin's maximum principle 및 Hamiltion-Jaccobi-Bellman equation 에 대한 간략한 소개", "문제 설정", "최적해의 존재성", "결론"이다.

2. 과제 수행(연구) 방법

우선 최적제어문제(optimal control problem, 이하 OC 문제)로 비용함수를 설정하고, 부등식 제한조건을 추가했다. OC 문제란 어떤 상태(state)와 이를 조절할 수 있는 제어항(control)이 있을 때, 비용함수를 최소화(또는 최대화)을 하는 제어항을 얻는 문제이다. 이를 수식으로 표현하면 다음과 같다.

$$\text{Min } J(x) = \int_a^b \Lambda(t, x(t), u(t)dt \text{ subject to } \frac{dx}{dt} = f\big(t, x(t), u(t)\big), t \in [a, b],\ \big(x(a), x(b)\big) = (A, B) \ldots (1)$$

이때 x(t)를 state trajectory, u(t)를 control trajectory 이다. 해당 식을 만족하는 u 는 2 가지 방법으로 구할 수 있다. 먼저 Pontryagin's maximum principle(이하 PMP)는 u_*가 (1)을 만족하면 이에 대한 필요조건을 다루는 정리이다. 본 연구에서는 이를 사용하지 않지만 최적해의 존재성에 대한 정리가 PMP 와 관련이 있어서 간략하게, 해당 principle 을 소개한다. PMP 를 수식으로 표현하면 다음과 같다(Clarke, 2013, [3]).

$$\mathrm{H(t, x, u, p, \eta)} = < \mathrm{p(t), f(t, x(t), u(t)} > -\eta * \Lambda(t, x(t), u(t) \ldots (2)$$
$$\mathrm{p'(t)} = -H_x(t,\ x_*, u_*, p) \ldots (3)$$
$$-p(b) \in N_E^L(x_*) \ldots (4)$$
$$H_u(t,\ x_*, u_*, p) = 0 \ldots (5)$$

(2)을 헤밀토니안, (3)을 adjoint equation, (4)를 transversality condition 이라고 한다. η는 0 과 1 중 하나의 값만을 취할 수 있는데, 0 은 제한조건이 강해서 cost 에 관계없이 제한조건을 만족하는 해가 유일하다는 것을 의미한다. $N_E^L(a)$ 은 limiting normals 이라고 하며, 이에 대한 자세한 설명은 [3]의 p240-p244 을 참조하면 된다. PMP 는 최적해가 만족해야 하는 필요조건을 다루고 있기 때문에, 해의 존재성과 유일성을 보장하지 않는다. 따라서 (2)-(4)를 만족하는 control 을 구해도, 그것의 실현가능성과 최적화여부를 별도로 증명해야 한다. 하지만 (5)을 통해 최적해가 될 수 있는 control 을 명시적(analytic)으로 표현할 수 있으며, 최적화 문제가 경계조건이 있는 편미분방정식으로 변환된다. 이로 인해 control 의 존재성과 유일성이 보장되면, 수치적으로 최적해를 쉽게 구할 수 있다.

그러한 문제에 대한 최적해의 존재성을 확인하기 위해, 정리 23.11(Francis Clarke, 2013, [3])과 정리 3.1(Hartl, R. F. and Sethi, Suresh and Vickson, Raymond, 1995)을 사용했다. 전자와 후자의 차이는 명시적으로 언급된 부등식 제한조건 유무이다.

정리 23.11(Francis Clarke, 2013): 다음과 같은 최적제어문제가 있다고 가정하자.

$$\text{Min } J(x, u) = \int_a^b \Lambda(t, x(t), u(t)dt$$

subject to $\frac{dx}{dt} = g_0(t,x(t)) + \Sigma g_i(t,x(t)) * |\mathrm{u(t)}|^j.\ (i = 1,2 \dots k)$

$$u(t) \in U(t)\ a.e, (t,x(t)) \in Q\ \ (x(a),x(b)) \in E$$

다음과 같은 조건들이 성립하면, 해당 문제는 최적해를 가질 수 있다.

(a)U(t)는 a.e 하게 closed and convex 하고, 집합 E 와 Q 는 closed 하다.

(b)state 의 초기값 x(a)에 대한 집합은 유계이다.

(c)피적분함수 Λ는 t, (x,u)에 대해 Lebesque measurable 하고, (x,u)에 대해 lower semi-continuous 하다.

(d)Λ는 모든 (t,x)에 대해 u 의 함수로써 convex 하고, bounded below 하다.

(e)$|\mathrm{u(t)}| \leq k(t)$ or $|\mathrm{u(t)}|^r + b \leq \Lambda(t,x(t),u(t)$을 만족하는 L^r((a,b)) 함수 k(t)나 상수 $r \geq 0, b$가 존재한다.

(f)함수 $g_i(i = 1,2 \dots k)$는 t 에 대한 measurable 하고 x 에 대해 연속인 함수이다. 그리고 모든 (t,x)∈Q에 대해, $|g_i(t,x)| \leq M(1+|x|)$을 만족하는 상수 M>0 이 존재한다.

정리 3.1(Hartl, R. F. and Sethi, Suresh and Vickson, Raymond, 1995): 다음과 같은 최적제어문제가 있다고 하자. 여기서 g, h, a, b 는 모두 벡터이며, 부등식이 성립한다는 것은 각 벡터의 성분들이 해당 부등식을 만족한다는 것을 의미한다.

$$\text{Min } J(x,u) = \int_a^b \Lambda(t,x(t),u(t)dt$$

$$\text{subject to } \frac{dx}{dt} = f(t,x(t),u(t),x(0) = x_0 \cdots \text{(a)}$$

$$g(t,x(t),u(t) \leq 0, h(t,x(t)) \leq 0 \dots (b)$$

$$a(b,x(b)) = 0, e(b,x(b)) \geq 0 \dots \text{(c)}$$

이때 제한조건 (c)의 LHS항들은 x에 대해 미분하면 선형독립이어야 하고, g는 서로 독립이어야 한다. 또한 함수 Λ, f, g, h, a, b는 정의역에 대해 연속이어야 한다. x 와 u 가 유계이고, 집합 N(x,t)={($\Lambda(t,x(t),u(t) + r, f(t,x(t),u(t))| r \leq 0, (b)\}$가 모든 (x,t)에 대해 convex 라고 하자. 그러면 제한조건 (a), (b), (c)를 만족하는 optimal solution (x_*, u_*)이 존재한다.

정리 3.1 은 cost function 에 라그랑지안 승수법을 사용하지 않았는데, 해당 방법은 state 들의 독립이어야 하기 때문이다. 이에 대한 증명이 implicit function theorem 으로 수행되었기 때문에(Clarke, 2013, [3], P175), control 이 state 에 영향을 주는 OC 문제에서는 라그랑지안을 사용할 수 없다. 또한 정리 3.1 의 N(x,t)는 f 에 의존하는데, f 가 비선형이면 N(x,t)에 대한 명제를 검토하기 힘들 수 있다.

(1)에 대한 최적해의 존재성은 HJB 방정식에 대한 해의 존재성과 동치다. 해당 방정식을 수식으로 표현하면 다음과 같다

$$\Lambda(t, x(t), u(t) = \phi_t\big(t, x(t), u(t)\big) + < \phi_x\big(t, x(t), u(t)\big), f(t, x(t), u(t) > \cdots (6)$$

(6)의 '='을 ≥으로 대체하고 등호는 항상 optimal solution (x_*, u_*)일 때만 성립한다고 하면, 이를 만족하는 함수를 ϕ 라고 한다. 이를 verification function 이라고 하며 ϕ로 가장 많이 사용되는 함수는 $V(t, B) = \int_a^t \Lambda(t, x(t), u(t)dt, (x(a), x(b) = (x_0, B)$ 이다. 이를 value function 이라고 한다. 이로 인해, $\phi = V(t, B)$일 때 (6)을 만족하는 해 (x(t),u(t))가 존재하면, 그것이 (1)의 최적해가 된다. 하지만 여기서는 (6)에 대한 해의 존재성을 확인하지 않았는데, 풀고자 하는 문제와 HJB 방정식과 충돌하기 때문이다. 이에 대해서는 차후 자세히 설명할 예정이다.

위의 정리들과 타겟을 추적하는 UV 에이전트에 대한 미분방정식을 종합하여, 효율적인 control 을 얻고자 한다. 그리한 최적해의 존재성을 정리 23.11 과 정리 3.1 을 통해 확인하고자 한다.

3. 과제 수행(연구) 결과

타겟의 위치는 상수이며, 이를 추적하는 UV 에이전트의 위치(q(t))와 속도(p(t))에 대한 미분방정식은 [1]의 정리 2.1 의 것과 같다(Sun-Ho Choi, Dohyun Kwon, Hyowon Seo, 2023). 다만 여기서는 단일 UV 에이전트를 사용했기 때문에, 여러개의 개체들과 관련된 term 은 제외되었다. 또한 타겟이 고정되었기 때문에, 그의 속도 $p_r(t)$는 항상 0 이 되낟. 그리고 q_r와는 타겟의 위치이며 $u(t)$는 UV 에이전트의 control vector 이다. 모든 벡터는 3 차원 유클리드 공간의 원소이다.

$$\frac{dq}{dt} = p(t) \ldots (7)$$

$$\frac{dp}{dt} = -\frac{\left||p(\mathrm{t})|\right|^2}{\left||q(\mathrm{t})|\right|^2} q(t) + c_q \left(\left||q(\mathrm{t})|\right|^2 q_r - < q(t), q_r(t) > q(t)\right) - c_p p(t) + u(t) \ldots (8)$$

다음과 같은 최적제어문제를 통해 효율적인 control u(t)를 설정하고자 한다. 이때 x(t)=(q(t),p(t))는 state trajectory, $f(t, x(t), u(t) = (\frac{dq}{dt}, \frac{dp}{dt})$ 은 미분방정식 제한조건이다. 또한 a은 상수이지만 b는 변수이며, 함수 f, k 와 state x 는 정의역에 대한 연속함수이다.

$$\mathrm{Min}\, J(x, u) = \int_a^b 1 + \left||p(\mathrm{t})|\right|^2 dt \ldots (9)$$

$$subject\ to\ \frac{dx}{dt} = f\big(t, x(t), u(t)\big) \ldots (10)$$

$$x(a) = (q_0, p_0), x(b) = (q_r(b), \alpha p_r(b)) \ldots (11)$$

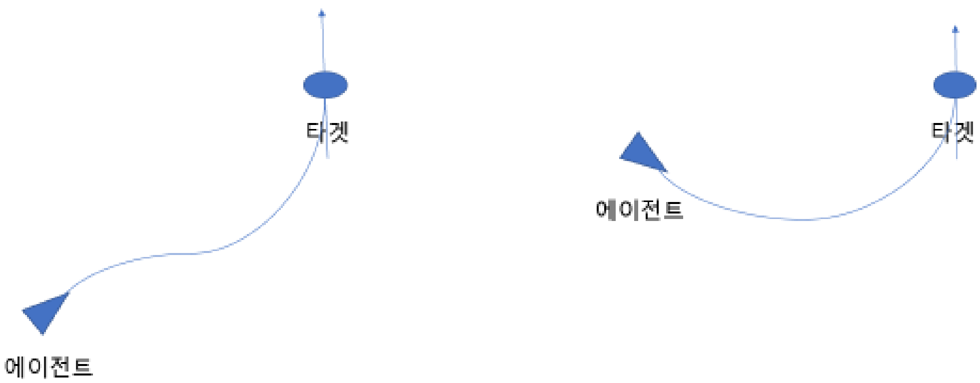
$$k\big(t, x(t), u(t)\big) = \left(\frac{\left\| cross\left(p(\mathrm{t}), \frac{dp}{dt}\right)\right\|^2}{\left\|p(\mathrm{t})\right\|^3} \right) \le \frac{1}{\left\|p(\mathrm{t})\right\|} \ldots (12)$$

(9)의 $\left\|p(\mathrm{t})\right\|^2$은 UV 에이전트의 이동거리를 계산하기 위한 항이고, 상수 1 은 이동시간을 최소화하기 위한 항이다. 이동거리 만을 최소화한 경로를 얻으면, 이를 따라가는 방법은 다양하다. 즉, 이동경로에 대한 reparametrization 의 결과가 무수히 많으며, 이로 인해 UV 에이전트의 속도가 매우 느릴 수 있다. 이러한 상황은 현실에서 부적절하기 때문에, 이동경로 뿐만 아니라 이동시간도 최소화하려고 한다.

속도의 정의로 인해, 이동시간의 최소화하는 control 은 최대 속도를 산출한다. 속도는 에이전트가 회전할 수 있는 각도와 반비례하기 때문에, 후자를 고려하여 경로를 설정해야 한다. 그렇지 않으면 UV 에이전트가 해당 경로를 이탈할 수 있다. 따라서 회전각도를 고려하여 제어항 u(t)를 설정해야 하는데, 이는 곡률로 표현될 수 있다. 곡률은 곡선의 휘어짐을 제는 척도인데, 이는 비행경로 각(flight path angle)과 기하학적으로 유사하다(Yao, Peng, Honglun Wang and Zikang Su, 2015, [4]). 이로 인해 부등식 제한조건 (12)을 설정했으며, 이는 $\left(\frac{\left\|cross\left(p(\mathrm{t}), \frac{dp}{dt}\right)\right\|^2}{\left\|p(\mathrm{t})\right\|^2} - 1 \right) \le 0$과 동치다. 여기서 cross(a,b)는 벡터 a 와 b 에 대한 외적이다.

final state 제한조건 (11)에서 $q(b) = q_r(b)$는 UV 에이전트가 타겟을 완전히 추적했음을 의미하고, $\mathrm{p}(\mathrm{b}) = \alpha p_r(b)$ 는 입사각 제한조건이다. 특정 임무에서는 UV 에이전트가 정해진 각도로 타겟에 진입하는 것이 중요한데, 우주정거장에 대한 도킹이나 전차의 취약점을 미사일로 공격하는 것이 있다. 타겟의 시야를 벡터 $p_r(b)$ 라고 하면, 입사각 제한조건은 cos(Θ)=k 가 된다(Θ는 $p_r(b)$ 와 $p(b)$ 간의 각도). 여기서 LHS 는 $\frac{<p_r(b), p(b)>}{\|p_r(b)\| * \|p(b)\|}$ 이며, 이는 $p(b) = k \frac{\| p(b) \|}{\|p_r(b)\|} p_r(b)$ 와 동치다. 여기서 $\alpha = k \frac{\| p(b) \|}{\|p_r(b)\|}$ 라고 하면, final state 제한조건 (11)를 얻을 수 있다.

제한조건 (11)의 final state 를 설정하면, UV 에이전트의 경로는 markov property 를 가지지 않는다. 즉, 현재 UV 에이전트의 위치는 미래의 위치에 영향을 준다. 이는 <그림 1>로 설명할 수 있는데, 오른쪽 UV 에이전트가 입사각 제한조건을 만족시킨다고 하면, 어느 시점에서 이동경로의 곡률이 왼쪽의 것보다 크다. 이로 인해 (9)에 대한 최적제어항을 HJB 방정식으로 구할 수 없는데, 해당 방정식은 principle of optimality 를 만족해야 하기 때문이다 이는 전체구간 [a,b]에서의 최적인 해 (x_*, u_*) 가 [a,c]에서도 최적임을 의미한다(a<c<b). 하지만 markov property 가 성립하지 않으면, principle of optimality 는 참이 아니다. 왜냐하면 principle of optimality 는 구간이 겹치지 않은 control 들이 서로 독립이라는 것을 의미하기 때문이다. 이로 인해, 본 연구에서는 HJB 방정식을 사용하지 않았다.



<그림 1>

미분방정식 (8)을 p(t)에 대해 2 번 편미분을 하면, $-\frac{2}{||q(t)||^2}q(t)$을 얻을 수 있다. 이는 p(t)가 state x(t)의 성분이고 행렬 및 벡터에 대한 미분연산으로 인해 성립한다. $-\frac{2}{||q(t)||^2}q(t)$의 모든 성분들이 항상 양수가 되지 않으며, 이는 각 성분들이 convex function 이 아니라는 것을 의미한다. N(x,t)가 (t,x(t))에 대해 convex set 이라고 하면, 이에 대한 사영도 convex set 이어야 한다. N(x,t)가 p(t)를 포함하는 공간으로 사영되었다고 하면, convex set 의 특성으로 인해 $f\left(t,x(t),u(t)\right)$의 $\frac{dp}{dt}$도 convex function 이어야 한다. 하지만 이는 $-\frac{2}{||q(t)||^2}q(t)$가 convex function 이 아니라는 사실과 모순된다. 따라서 N(x,t)는 (t,x(t))에 대해 convex set 이 아니다. 따라서 정리 3.2 으로 OC 문제 (9)-(12)가 최적해를 가진고 주장할 수 없다.

제약조건 (12)는 정리 23.11 의 control set U(t)를 설정하는데 사용되었다. 이로 인해, 기존 OC 문제에서 부등식 제한조건이 소멸하여, 정리 23.11 의 가정을 검토해볼 수 있다. 정리 3.1 의 (a)가 성립하지 않을 수 있음을 통해, 해당 정리로 최적해의 존재성을 증명할 수 없음을 보일 것이다. 변환된 부등식 $\left(\frac{\left\|cross\left(p(t),\frac{dp}{dt}\right)\right\|^2}{||p(t)||^2}-1\right)\leq 0$에서 $s(t)=p(t)/\|p(t)\|$으로 p(t)를 정규화한 다음에, $\frac{dp}{dt}$와 삼각부등식을 사용하여 다른 부등식을 얻을 것이다. 이를 수식으로 표현하면 다음과 같다.

$$\left(\frac{\left\|cross\left(p(\mathrm{t}),\frac{dp}{dt}\right)\right\|^2}{\|p(\mathrm{t})\|^2}-1\right)\leq 0 => \left\|cross\left(s(\mathrm{t}),\frac{dp}{dt}\right)\right\|^2 \leq 1$$

$$=> \left\|cross\left(s(t),-\frac{\|p(\mathrm{t})\|^2}{\|q(\mathrm{t})\|^2}q(t)+c_q\left(\|q(\mathrm{t})\|^2 q_r - < q(t),q_r(t) > q(t)\right)-c_p p(t)+u(t)\right)\right\| \leq 1$$

$$=> \left\|cross\left(s(\mathrm{t}),u(t)\right)\right\| \leq 1+\left\|cross\left(s(t),-\frac{\|p(\mathrm{t})\|^2}{\|q(\mathrm{t})\|^2}q(t)+c_q\left(\|q(\mathrm{t})\|^2 q_r - < q(t),q_r(t) > q(t)\right)-c_p p(t)\right)\right\| \ldots(13)$$

$$=> \left\|\left(u(t)\right)\right\| \leq 1+a(t)\left\|cross\left(s(t),-\frac{\|p(\mathrm{t})\|^2}{\|q(\mathrm{t})\|^2}q(t)+c_q\left(\|q(\mathrm{t})\|^2 q_r - < q(t),q_r(t) > q(t)\right)-c_p p(t)\right)\right\| \ldots(14)$$

이때 $a(t)$는 $\left\|\left(s(t)\right)\right\|\sin\left(\mathrm{w(t)}\right)$이며, w(t)는 u(t)와 s(t)간의 각도이다. 등식 (13)과 (14)의 RHS 는 시간 t 에 대한 함수이며, 이는 control u(t)의 경계가 시간에 따라 변한다는 것을 의미한다. 따라서 control set U(t)의 convex 여부는 보장할 수 없다. 또한 이와 같은 움직이는 경계(moving boundary)에서 미분방정식의 해가 존재한다는 것을 증명하기 어렵다. 이로 인해, U(t)가 convex set 이어도 PMP 의 adjoint equation (3)을 해결하기 쉽지 않다

4.결론 및 제언

정리하면 미분방정식 (8)이 state (q(t),p(t))에 대해 convex 하지 않고, 부등식 제한조건 (12)를 (14)로 변환하여 control u(t)의 움직이는 경계를 얻을 수 있다. 이로 인해, 정리 3.1 과 정리 23.11 을 통해 optimal control 의 존재성을 증명할 수 없다. 해의 존재성을 위해 일부 식을 변형하거나, 제한조건을 제거할 수 있다. 하지만 여기서는 그러한 방법을 적용하기 어려운데, 우선 미분방정식을 변형하는 것은 타겟에 대한 UV 에이전트의 수렴성을 보장하지 않는다(Sun-Ho Choi, Dohyun Kwon, Hyowon Seo, 2023).

또한 곡률 k(t,x(t),u(t))은 UV 에이전트의 속도 p(t)에 의존하기 때문에, control u(t)의 경계는 시간에 대한 함수이다. 그러한 제한조건을 제거하면 optimal control 이 존재할 수 없다. 에이전트의 초기 위치와 타겟 간의 직선으로 이동하다가, final state 근방에서 입사각 제한조건을 만족시키기 위해 급격하게 꺾는 것이 최적의 의 경로이기 때문이다. 그리고 꺾기 시작한 위치와 타겟 간의 거리가 작을수록, 경로는 보다 최적이 된다. 그러나 이는 현실적으로 구현할 수 없다. 따라서 곡률에 대한 제한조건을 제거할 수 없다.

따라서 휴리스틱 알고리즘, 유전 알고리즘이나 강화학습 등을 통해, 보다 효율적인 경로를 설정해야 한다. 이들은 전역적으로 최적인 해를 산출하지 않을 수 있지만, 논문 [1]의 control 보다 효율적일 수 있다. 또한 그러한 알고리즘들은 제한조건을 반영한 control 을 보다 쉽게 얻을 수 있다. 다만 그러한 수치해는 타겟에 대한 UV 에이전트의 수렴성을 분석하는데 어려움을 야기한다.

5. 참고문헌 및 조사기관

[1]Sun-Ho Choi, Dohyun Kwon, Hyowon Seo, Multi-agent system for target tracking on a sphere and its asymptotic behavior, Communications in Nonlinear Science and Numerical Simulation, Volume 117, 2023

[2]Hartl, R. F. and Sethi, Suresh and Vickson, Raymond, A Survey of the Maximum Principles for Optimal Control Problems with State Constraints. SIAM Review, Vol. 37,No. 2, pp. 181-218, 1995

[3]Francis Clarke, Functional Analysis, Calculus of Variations and Optimal Control, 2013

[4]Yao, Peng, Honglun Wang and Zikang Su. "Real-time path planning of unmanned aerial vehicle for target tracking and obstacle avoidance in complex dynamic environment." Aerospace Science and Technology 47 (2015): 269-279